

## Utilisation des bases de données bibliographiques spécialisées en traduction médicale

Sylvie Vandaele

Volume 46, numéro 1, mars 2001

Traduction médicale et documentation / Medical translation and documentation

URI : <https://id.erudit.org/iderudit/004534ar>

DOI : <https://doi.org/10.7202/004534ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Vandaele, S. (2001). Utilisation des bases de données bibliographiques spécialisées en traduction médicale. *Meta*, 46(1), 103–116.  
<https://doi.org/10.7202/004534ar>

Résumé de l'article

Les bases de données bibliographiques spécialisées constituent une riche source d'informations à la fois terminologiques et notionnelles. L'élaboration de stratégies spécifiques permet au traducteur de résoudre différents problèmes rencontrés dans le cours de sa pratique professionnelle : constitution d'une bibliographie spécialisée ; éclaircissement de nouveaux concepts ; recherches terminologiques (notamment en ce qui a trait à l'évolution du sens et à la néologie) ; vérification de données plus ou moins bien exprimées dans le texte à traduire. Différents éléments permettent de réfléchir à la validité des éléments colligés lors des recherches : par exemple, diversité du corpus, qualité de la source, origine de l'article et des auteurs, diachronie et synchronie, écoles de pensée.

# Utilisation des bases de données bibliographiques spécialisées en traduction médicale

SYLVIE VANDAELE

*Université de Montréal, Montréal, Canada*

## RÉSUMÉ

Les bases de données bibliographiques spécialisées constituent une riche source d'informations à la fois terminologiques et notionnelles. L'élaboration de stratégies spécifiques permet au traducteur de résoudre différents problèmes rencontrés dans le cours de sa pratique professionnelle: constitution d'une bibliographie spécialisée; éclaircissement de nouveaux concepts; recherches terminologiques (notamment en ce qui a trait à l'évolution du sens et à la néologie); vérification de données plus ou moins bien exprimées dans le texte à traduire. Différents éléments permettent de réfléchir à la validité des éléments colligés lors des recherches: par exemple, diversité du corpus, qualité de la source, origine de l'article et des auteurs, diachronie et synchronie, écoles de pensée.

## ABSTRACT

Specialized bibliographic databases represent an important source of information, both at terminological and conceptual levels. Specific strategies allow translators to solve a number of difficulties met during professional practice: building up of specialized bibliographies; understanding of new concepts; terminology searches, particularly related to changes of meaning and to neology; control of poorly expressed data in the source text. The validity of information being collected may be evaluated according to a number of criteria: corpus diversity; origin of publication and authors; diachrony and synchrony; schools of thought.

## MOTS-CLÉS/KEYWORDS

Bases de données, informations terminologiques, informations notionnelles, élaboration de stratégies, recherches terminologiques

Les bases de données bibliographiques sont offertes, depuis quelques années, sur support électronique. Elles constituent un outil privilégié pour le chercheur et le professionnel d'un domaine qui veut garder ses connaissances à jour. Mais tant le traducteur que le terminologue y trouvent aussi leur compte: pour peu qu'ils aient accès à Internet ou à une bibliothèque qui offre un accès sur place ou par modem aux bases auxquelles elle est abonnée, ils ont au bout des doigts des corpus qui leur permettent, en un temps record, de mener des recherches qui auraient pris auparavant bien de leur temps. Mais toute médaille a son revers: l'efficacité est tributaire de la maîtrise de l'outil. En biomédecine, il existe plusieurs bases, chacune ayant sa spécificité. Afin de guider les professionnels langagiers dans les dédales électroniques des bases de données bibliographiques, nous présenterons tout d'abord différentes bases utiles dans le domaine, puis certaines stratégies de recherche visant à résoudre des problèmes spécifiques à la traduction biomédicale.

### Contenu

Les bases de données bibliographiques répertorient dans des fiches (électroniques) les références d'articles ou d'ouvrages publiés dans différents domaines. Chaque fiche correspond à une référence et comporte différents champs, au moins le nom des auteurs, l'adresse de ces derniers, le titre du document, ainsi que la source. S'y ajoutent, selon les bases et les années, un numéro d'identification, des mots-clés ou des descripteurs autorisant la recherche par sujet, le résumé (le plus souvent en anglais), parfois le titre dans la langue d'origine, le numéro d'ISSN, le type de publication, la date d'inscription dans la base, etc.

Certaines bases contiennent des articles complets, soit ponctuellement (par exemple, les articles traitant de médecine factuelle [*Evidence-Based Medicine Reviews*] dans Medline, version Ovid), soit parce que la nature même de la base le veut ainsi (Cochrane Database). La plupart des références indexées dans les bases sont en anglais, ce qui reflète le fait que la vaste majorité des articles de biomédecine sont publiés dans cette langue, même par les auteurs non anglophones. Par exemple, dans Pascal, base réalisée en France, 76 % des références sont en anglais, 9,6 %, en français, 6 %, en russe, 5,1 %, en allemand, 3,3 %, dans d'autres langues (espagnol, italien, japonais, polonais...). Toutefois, les bases indexent, à des degrés divers, des périodiques et des ouvrages écrits dans d'autres langues que celle de Shakespeare. Certaines ne rendent accessibles que la version anglaise du résumé (Medline, Current Contents) et le titre dans la langue d'origine (Medline), tandis que d'autres permettent aussi de consulter le résumé dans la langue d'origine, notamment le français (Pascal). Aussi, même si la documentation dans les langues autres que l'anglais reste minoritaire, les bases représentent une voie d'accès privilégiée à ce corpus.

Chaque base cible un ensemble de sous-domaines biomédicaux particuliers. La liste que nous présentons n'est sans doute pas exhaustive, mais elle autorise l'accès à un vaste éventail de connaissances qui recouvre l'ensemble de la biomédecine. Il faut savoir aussi que certains périodiques sont indexés dans plusieurs bases, ce qui signifie qu'une certaine redondance peut être observée si l'on est amené à en consulter plus d'une.

Un premier type de bases de données collige essentiellement des références de périodiques et d'autres documents tels que des monographies ou des thèses. La plus connue est **Medline**. Elle fait partie de MEDLARS (*MEDical Literature Analysis and Retrieval Systems*), qui est un ensemble informatisé de plus de quarante bases offertes par la *National Library of Medicine* (NLM). Medline contient plus de 10 millions de références et 30 000 s'y ajoutent chaque mois. On y trouve des revues consacrées aux différentes spécialités médicales, aux essais cliniques et à certaines sciences connexes (sciences biologiques et physiques, sciences humaines, sciences de l'information, dans la mesure où elles concernent la médecine et les soins de santé). Certains périodiques en français y sont indexés, notamment *Thérapie*, dans lequel on trouve de nombreux articles consacrés à des essais cliniques, et la *Revue du praticien-monographie*. Par contre, *M/S Médecine Sciences*, revue de haut niveau particulièrement utile pour suivre la recherche biomédicale de pointe réalisée par des chercheurs francophones (essentiellement France et Québec), n'y est pas indexée. Medline indexe également de nombreux titres relatifs à l'odontologie et aux sciences infirmières. Toutefois, ces dernières sont la spécialité de **CINAHL** (*Cumulative Nursing & Allied*

*Health*), qui se préoccupe également des techniques cardiopulmonaires, des urgences, de l'éducation en santé, des techniques de laboratoire, de l'ergothérapie... **Current Contents** regroupe par modules consultables globalement (CCALL) ou indépendamment les périodiques relevant de la médecine clinique (CLIN), de l'agriculture (AGRI), des arts et sciences humaines (ARTS), des sciences sociales et comportementales (BEHA), des sciences de la vie (LIFE), de la physique-chimie et des sciences de la terre (PHYS), du génie et de l'informatique (TECH). Selon le type d'abonnement auquel il est possible d'avoir accès, on pourra consulter tous les modules ou seulement certains. En biomédecine, on s'orientera vers les modules CLIN et LIFE. Contrairement à Medline, Current Contents indexe la revue *M/S Médecine Sciences*<sup>1</sup>. **Biological abstracts** concerne l'ensemble des sciences de la vie, notamment la biologie et la biomédecine (botanique, zoologie, microbiologie, médecine clinique, vétérinaire et expérimentale, biotechnologie, environnement et agriculture). **PsycINFO** est centré sur la psychologie et les sciences connexes (médecine, psychiatrie, sciences infirmières, sociologie, éducation, pharmacologie, physiologie, linguistique...). **CAB Abstracts** (*Commonwealth Agricultural Bureau Abstracts*) concerne la médecine vétérinaire et les domaines apparentés (agriculture, nutrition, biotechnologie, foresterie, etc.). **Pascal** est une base de données multidisciplinaire réalisée en France conjointement par l'Institut de l'information scientifique et technique (INIST) et le Centre national de la recherche scientifique (CNRS) signalant 13 millions de références en sciences, technologie et médecine depuis 1973. Elle permet d'accéder à de nombreuses revues écrites en français. Certaines bases se spécialisent sur des sujets brûlants, comme par exemple **BIOETHICSLINE**, qui porte sur les questions d'éthique et de déontologie en santé<sup>2</sup>. Citons également deux catalogues grâce auxquels il est possible de commander en ligne les documents moyennant une inscription. D'une part, **ARTICLE@INIST**, qui est réalisé par le CNRS et l'INIST, donne accès gratuitement à des catalogues de monographies (thèses, rapports de congrès), de revues, de numéros de revues et d'articles. Fait intéressant, les résumés des revues francophones sont accessibles dans la langue de Molière (ce qui n'est pas le cas dans Medline, par exemple, où seuls apparaissent les résumés en anglais). L'obtention du mot de passe permet de consulter les résumés et de commander en ligne des documents. Le Canada n'est pas en reste puisqu'il propose le catalogue de l'**ICIST** (Institut canadien de l'information scientifique et technique). Anciennement SwetScan, le catalogue de l'ICIST collige des documents relatifs aux sciences physiques et de la vie, de l'ingénierie, de la technologie et des sciences de la santé. La collection des documents médicaux est très importante, la spécialité de l'ICIST étant notamment les comptes rendus de conférences médicales. L'inscription est requise pour la plupart des services offerts, notamment la commande de documents en ligne.

Un deuxième type de bases regroupe des articles plein texte: **Evidence Based Medicine Reviews** (*ACP Journal Club's Best Evidence*) propose des examens critiques d'articles sélectionnés dans plus de 90 périodiques portant sur des études menées en médecine interne et en médecine spécialisée. **Cochane Database of Systematic Reviews**, quant à elle, présente des revues d'essais cliniques qui sont préparées par la *Cochrane Collaboration*, organisme international à but non lucratif dont l'objectif est de synthétiser les résultats provenant de divers essais cliniques.

### Périodes couvertes et ampleur

Le tableau 1 présente les principales bases évoquées ci-dessus, avec les périodes couvertes et le nombre de périodiques recensés par chacune d'entre elle. Les périodes couvertes sont variables. La base la plus ancienne, Medline, existe depuis presque 40 ans.

**Tableau 1 : Périodes couvertes et ampleur des principales bases de données bibliographiques**

Base	Période	Ampleur de la base
Medline	1966 -	3 900 périodiques
BIOETHICSLINE	1973 -	> 53 000 fiches
Biological Abstracts	1980 -	> 6 500 périodiques
CINAHL	1982 -	> 1 200 périodiques
Current Contents	1994 -	6 600 périodiques
CAB Abstracts	1973 -	> 11 000 périodiques
Biosys Previews	1989 -	6 500 périodiques
PsycINFO	1987 -	> 1 300 périodiques
Pascal	1973 -	Thèses, monographies, brevets, rapports (2 %), congrès (14 %), périodiques (> 5 000, dont 40 % France et Europe: 84 %)
ARTICLE@INIST	1990 -pour la base « Articles »	> 25 000 périodiques > 700 000 numéros de revues > 6 000 000 articles > 160 000 monographies
ICIST	1993 - ...	14 000 périodiques
Evidence Based Medicine Reviews	1991-2000	> 90 périodiques
Cochrane Database of Systematics Reviews	Mise à jour permanente	—

### Comment y accéder ?

De nombreuses bibliothèques, notamment celles des universités, sont abonnées à des bases de données, bibliographiques ou non. Par exemple, à l'Université de Montréal (<[www.bib.umontreal.ca](http://www.bib.umontreal.ca)>), le serveur Hermès offre, sur place ou via Internet, Evidence Based Medicine Reviews, Biological abstracts, CAB Abstracts, CINAHL, Cochrane Database, Current Contents, Medline, PsycINFO et Evidence Based Medicine Reviews dans le format Ovid. Pour tous ceux qui n'ont pas accès aux services universitaires, Medline est accessible gratuitement sur Internet par l'intermédiaire de la NLM (*National Library of Medicine*): d'une part, la nouvelle version de PubMed (<[www.ncbi.nlm.nih.gov/entrez](http://www.ncbi.nlm.nih.gov/entrez)>), qui offre aussi l'accès à Premedline (base dans laquelle apparaissent transitoirement les références avant leur traitement et leur indexation dans Medline), et à des bases non bibliographiques (Nucleotide, Protein, Genome, Structure et Popset); d'autre part, Internet Grateful Med<sup>3</sup> (<[igm.nlm.nih.gov](http://igm.nlm.nih.gov)>), qui donne également accès à une quinzaine de bases (Healthstar, Aidsdrugs, etc.). Une autre solution consiste à s'inscrire à un fournisseur comme Healthgate (<[www.healthgate.com](http://www.healthgate.com)>). Ce site permet de consulter Medline gratuitement mais avec un format d'interrogation réduit. D'après une revue comparative parue dans *Le Clinicien* (Beaudet et Hébert 1998), le format Ovid offre le plus com-

plet des modes d'interrogation de Medline et autorise des stratégies de recherche à la fois plus sophistiquées et plus conviviales que celles des versions gratuites. C'est également ce que nous avons nous-même constaté<sup>4</sup>.

Mis à part les bases offertes par la *National Library of Medicine*, l'accès à la plupart des bases autres que Medline est payant et nécessite, si l'on ne peut avoir accès à une bibliothèque abonnée, d'avoir recours à un fournisseur qui pourra en proposer plusieurs. Ainsi, Healthgate offre notamment l'accès à CINAHL, PsycInfo, EMBASE (laquelle regroupe les références en fonction de treize spécialités, telles que anesthésiologie, cardiologie, etc.), ainsi qu'à des bases spécialisées telles que AgeLine (gériatrie), AIDSDRUG, AIDSTRIAL et AIDSLINE (sida), CancerLit (oncologie), etc. Healthgate autorise la consultation des titres, mais l'affichage du résumé est payant.

Quant à Pascal, lorsqu'il n'est pas possible d'y avoir accès par un serveur universitaire, l'accès est payant (<[www.inist.pascal/pascal.htm](http://www.inist.pascal/pascal.htm)>). Toutefois, il pourrait peut-être devenir gratuit, comme l'y incitait le rapport du député français Bloche (cité dans Cassagne, Darmoni *et al.* 2000 : 4). Lorsque plusieurs bases sont accessibles grâce à un outil de consultation intégré, comme c'est le cas à l'Université de Montréal (version Ovid), la même stratégie de recherche peut être sauvegardée et directement réutilisée pour consulter une autre base. C'est un facteur dont il faut tenir compte quand il s'agit de choisir un mode d'accès.

## Stratégies de recherche

### a. Documentation thématique

L'objectif visé est de constituer une documentation sur un sujet donné. Pour cela, on tirera parti de la présence d'un thésaurus de descripteurs utilisés pour indexer les références et pouvant être organisés en arborescence (Medline, CINAHL) ou non (PsycINFO). Current Contents fait appel à une liste de mots-clés générés à partir du titre de la référence, mais qui ne sont pas organisés en arborescence.

Medline fait appel à un ensemble de descripteurs organisés en arborescence conceptuelle, MeSH (*Medical Subject Headings*). Le vocabulaire MeSH comporte environ 45 000 termes. Il est utilisé pour indexer manuellement les articles et fait l'objet d'une mise à jour permanente<sup>5</sup>. CINAHL et PsycINFO ont tous deux un thésaurus : pour la première, *NISO Thesaurus Database*, qui est identique au vocabulaire MeSH de Medline à 70 %, mais qui comporte 2000 termes spécifiques ; pour la deuxième, l'*American Psychological Association's Thesaurus Psychological Index Terms*. Dans chacune de ces bases en format Ovid, les mots utilisés pour la recherche sont automatiquement mis en correspondance avec les termes du thésaurus.

Ainsi, si on lance une recherche dans Medline (format Ovid) avec le terme « liver », celui-ci est automatiquement mis en correspondance (*mapping*) avec une liste de descripteurs (*subject headings*) parmi lesquels il sera possible de choisir celui qui se rapproche le plus du thème recherché. Cette fonction n'est pas nécessairement accessible dans les différentes versions de Medline offertes sur Internet. Le descripteur choisi apparaît alors au sein d'une arborescence conceptuelle, ce qui permet d'élargir ou de restreindre la recherche (Tableau 2). Il est ensuite possible de préciser cette dernière par le choix de l'un des 80 qualificatifs (*MeSH subheadings*) apparaissant sur des écrans successifs (par exemple, « Diagnosis », « Economics », etc.). Les qualificatifs

sont apposés aux descripteurs MeSH pour en préciser le sens selon la syntaxe *heading/subheading*, par exemple, «liver diseases/diagnosis», ce qui revient à restreindre la recherche. Dans cet exemple, ce sont les références traitant du diagnostic des hépatopathies qui seront sélectionnées.

La fonction d'explosion (*explode*) permet de rechercher les documents indexés avec un terme *et* tous les termes situés «en-dessous» de lui dans l'arborescence.

**Tableau 2 : Extraits de la liste des descripteurs et de l'arborescence.**

La sélection de *liver diseases* dans la liste des descripteurs fait apparaître l'arborescence conceptuelle dans laquelle il s'insère. Il est possible alors de préciser la recherche (par exemple, sélectionner *Fatty Liver*) ou de l'élargir (remonter à *Digestive System Diseases*).

Liste des descripteurs	Arborescence
Liver cirrhosis, biliary Liver cirrhosis, experimental <b>Liver diseases</b> Liver diseases, alcoholic ...	Digestive System Diseases <b>Liver Diseases</b> Cholestosis, Intrahepatic Fatty Liver Focal Nodular Hyperplasia ...

Par ailleurs, les descripteurs MeSH sont associés à des annotations (*Scope Note*) comprenant une définition, des notes, l'année d'indexation, des références, des renvois à d'autres mots-clés, etc. (Tableau 3)<sup>6</sup>.

**Tableau 3 : Annotations associées aux descripteurs de Medline dans la version Ovid**

Scope Note – Proteome
MeSH HEADING: PROTEOME SCOPE: The protein complement of an organism coded for by its genome. NOTE: coord with precoord organism protein + organism YEAR OF ENTRY: 2000 REFERENCES: See related: GENOME Used For: Proteomes

Le recours aux descripteurs constitue un mode de recherche précieux quand l'objectif est de constituer une documentation thématique. L'organisation des descripteurs en arborescence (Medline, CINAHL) est particulièrement utile, car il est ainsi possible de restreindre la recherche (bien qu'il existe d'autres stratégies pour restreindre les recherches, notamment l'utilisation des opérateurs booléens) ou au contraire de l'élargir. Enfin, notons que la fonction de mise en correspondance n'existe pas dans Current Contents, le résultat de la recherche par mot-clé dépendant simplement de la présence ou de l'absence de celui-ci dans la liste.

### **b. Recherches terminologiques ponctuelles**

Plusieurs stratégies peuvent être utilisées pour mener à bien des recherches terminologiques ponctuelles.

La première tire parti des annotations associées aux descripteurs du thésaurus de Medline : en effet, elles contiennent des informations de type définitoire (Tableau 3). C'est l'une des premières ressources à consulter lorsqu'on rencontre des néologismes encore absents des dictionnaires.

La deuxième stratégie consiste à rechercher des contextes explicatifs ou définitoires au sein même des résumés d'articles. On fera appel, dans ce cas, à la fonction de recherche de « mot libre » (*text word* sous format Ovid), qui permet de rechercher les mots essentiellement dans le titre et le résumé. Selon les résultats obtenus, on pourra étendre ou restreindre les recherches à l'aide de différents critères (voir plus bas). On notera un inconvénient, toutefois : les bases présentées ne permettent pas, actuellement, de combiner recherche de termes et recherche de patrons explicatifs ou définitoires. Si l'on souhaite faire appel à ce type de technique, il faudra sauvegarder les références sous un format texte (voir « Constitution de corpus » ci-dessous) et avoir recours à des outils logiciels indépendants.

Prenons comme exemple les termes suivants, récemment apparus : *proteom*, *proteomic*, *proteomics*. Les interrogations ont été menées dans la version Ovid de Medline en septembre 2000, pour la période 1997-2000. Dans le mode d'interrogation par mot-clé, les trois termes sont automatiquement reliés au même descripteur « *proteome* » et aux annotations correspondantes (Tableau 4). Notons que 117 références sont ainsi trouvées (dont aucune en français) et qu'il n'est pas certain que le terme apparaisse dans le titre ou le résumé. La première référence indexée date de 1999, la troisième comprend un contexte définitoire<sup>7</sup>. Si l'on utilise la fonction de recherche plein texte (terme dans le titre ou le résumé), on obtient 68 références pour *proteomic*, 216 pour *proteome* et 122 pour *proteomics*. Si l'on combine les trois recherches avec l'opérateur OR, on aboutit à un total de 345 références, ce qui signifie que les trois ensembles de références se recoupent dans une certaine mesure ( $345 < 68 + 216 + 122$ ), tout simplement parce que les différents termes coexistent au sein de plusieurs résumés ou titres. Parmi ces références, une seule est rédigée en français. La référence la plus lointaine date de 1997, la deuxième comportant un contexte définitoire et explicatif<sup>8</sup>.

**Tableau 4 : Résultats obtenus avec deux stratégies d'interrogation différentes de Medline (version Ovid), l'interrogation par mot-clé et l'interrogation plein texte**

Requête	Résultats (nombre de références)
– par mot-clé	La mise en correspondance amène à un résultat identique (arborescence des descripteurs) :
<i>Proteome</i>	<i>Proteins</i> 18 609
<i>Proteomics</i>	<i>Proteome</i> 117
<i>Proteomic</i>	
– plein texte	Les résultats sont différents :
<i>Proteome</i>	216
<i>Proteomics</i>	122
<i>Proteomic</i>	68
<i>Proteome OR proteomics OR proteomic</i>	345



Ainsi, le résultat des recherches dépend de la stratégie utilisée. Deux astuces permettent d'accélérer la recherche de contextes définitoires ou explicatifs notamment dans le cas des néologismes. La première consiste à combiner les résultats de la recherche par mots-clés et ceux de la recherche plein texte. On sélectionnera ainsi des références qui portent sur les notions recherchées *et* dont le résumé ou le titre comportent le terme correspondant, ce qui peut être très utile pour élargir la recherche (opérateur OR) ou la restreindre (opérateur AND). La deuxième consiste à consulter les références en commençant par les plus anciennes. En effet, on retrouve assez facilement des contextes explicatifs ou définitoires dans les premiers articles dans lesquels le terme apparaît. Si la recherche est infructueuse, il sera toujours temps d'élargir la consultation.

### *c. Constitution de corpus*

Les corpus sont particulièrement utiles quand vient le temps d'étudier la phraséologie et les collocations. Or, les bases de données constituent en elles-mêmes des corpus de textes dont les caractéristiques ont l'avantage d'être bien définies. Tout d'abord, elles sont représentatives de l'usage accepté d'un domaine. En effet, les articles indexés dans les bases proviennent de périodiques qui soumettent leurs publications à un examen dit « par les pairs » (*peer reviewed articles*). Cela signifie que des experts d'un domaine faisant partie d'un comité de lecture se prononcent sur l'acceptabilité scientifique d'un document. Ce processus vise à garantir la qualité des travaux publiés. Par conséquent, tant sur le plan notionnel que terminologique, la valeur de ces sources est de loin supérieure à celle de la plupart des textes que l'on retrouve sur Internet, dont il est parfois impossible de juger la qualité. En raison du processus d'examen des articles, on peut en effet penser que les différences d'usage ou les modifications du sens d'un terme sont davantage le reflet de l'évolution globale du domaine ou d'écoles de pensée différentes que le fait isolé d'un auteur dont il n'est jamais exclu qu'il soit dans l'erreur. C'est d'ailleurs un point extrêmement important dans le cadre de recherches terminologiques ponctuelles, notamment pour les traducteurs débutants qui en sont au stade de la construction d'un bagage cognitif de fond et qui n'ont pas encore suffisamment de données en main pour porter avec assurance un jugement sur les sources.

Ensuite, la densité des termes dans un corpus provenant d'une base de données bibliographiques est très élevée, car il est possible d'extraire les données en les centrant sur des termes spécifiques. À titre d'exemple, la fréquence du syntagme *second messenger* dans un corpus extrait de Medline entre 1966 et 1999 (plus de 4000 occurrences pour un corpus de 1 570 000 mots converti en format texte et téléchargé) est environ 400 fois plus importante que dans un ouvrage de référence tel que *Molecular Cell Biology* (Lodish *et al.* 2000). De plus, un grand nombre d'auteurs sont représentés dans le corpus ainsi constitué. Enfin, les fonctions permettant de limiter ou de combiner les recherches (voir plus bas) permettent, selon les besoins, de restreindre le corpus à un sous-domaine ou à une période donnée, ou au contraire de l'élargir si les réponses obtenues ne sont pas satisfaisantes, notamment sur le plan statistique.

La sauvegarde des références (qui sont envoyées en format texte par courrier électronique), y compris le résumé, permettra de conserver des corpus qui pourront être exploités à l'aide soit de concordanciers (Naturel, Le concordeur...) soit de logiciels de bibliographie spécialisés (Endnote, Procite, Reference Manager...).

Enfin, il ne faut pas oublier que Medline, au moins dans sa version Ovid, offre l'accès plein texte à environ 80 périodiques et que certaines bases, par essence, sont constituées de textes complets, comme *Evidence Based Medicine Reviews* et *Cochrane Database*.

#### ***d. Limitations et combinaisons des recherches***

Que faire lorsque le nombre d'articles retrouvés est trop important ? Toutes les bases fournissent, à des degrés divers, la possibilité de restreindre les recherches. La version OVID de Medline est à cet égard l'une des plus complètes. Il est possible de limiter les recherches en fonction de nombreux critères (Tableau 5).

**Tableau 5 : Critères permettant de limiter les recherches dans la version Ovid de Medline. Les points de suspension indiquent que des sous-catégories sont accessibles.**

Evidence Based Medicine Reviews	Male
Topic Reviews <Cochrane>	Female
Article Reviews <Best Evidence>	AIM Journals
Human	Languages...
English Language	Age Groups...
Review Articles	Publication Year...
Abstracts	Publication Types...
Latest Update	Animal Types
Animal	Journal Subsets...

Les articles de synthèse (*review articles*) sont particulièrement utiles pour prendre connaissance des dernières avancées dans un domaine de recherche particulier. Ils complètent très bien les ouvrages de référence qui ont environ cinq ans de retard sur l'état actuel des recherches. En fait, lorsque l'on effectue une recherche notionnelle, il est conseillé de ne consulter les articles très spécialisés que pour répondre à des questions bien précises.

La sélection de la langue (*Languages...*) dans laquelle le corps de l'article est écrit est particulièrement utile en traduction. On pourra, à l'aide de ce mode de restriction, retrouver la documentation rédigée dans la langue d'arrivée. La plupart des versions de Medline accessibles sur Internet permettent de sélectionner l'anglais ou le français. Là encore, la version Ovid est particulièrement performante, puisque plus de 60 langues sont représentées. Une fois la référence en main, il suffit de se procurer l'article, soit directement en ligne, soit dans une bibliothèque. De plus, toujours dans la version Ovid, le titre de l'article apparaît dans la langue d'origine. L'examen des titres d'article peut alors fournir de précieux indices quant à l'usage de certains termes, bien que cela soit insuffisant dans certains cas (par exemple : faible fréquence, terme apparu trop récemment, contexte non significatif). Enfin, certaines bases (Pascal) permettent de consulter directement les résumés en français.

La limitation des recherches permet également de poursuivre différents objectifs spécifiques. Il peut être parfois très important de prendre en considération le pays d'origine de l'auteur, car le choix d'une terminologie, même en anglais, peut refléter l'influence prédominante d'une école de pensée nord-américaine ou européenne. C'est notamment le cas pour l'utilisation, dans le contexte de l'anesthésie, des termes *hypnosis* et *narcosis* (Vandaele 2001).

Le choix d'une plage temporelle, soit d'une seule année, soit de plusieurs années permet d'avoir un aperçu de la fréquence d'usage d'un terme en fonction du temps, de déterminer le moment auquel il est apparu et d'examiner l'évolution du sens au cours du temps.

Fait intéressant, la plupart des bases offrent la possibilité d'avoir recours à des opérateurs booléens tels que AND, OR et NOT. L'opérateur AND est restrictif, puisqu'il impose que les critères choisis soient tous respectés. L'opérateur OR impose qu'au moins un des critères soit respectés. Quant à NOT, il permet d'exclure certaines données. On pourra consulter la fonction d'aide de chacune des bases consultées pour connaître les modes d'utilisation particuliers offerts par chacune d'entre elles.

La maîtrise progressive des modes de limitation des recherches autorise l'élaboration de stratégies de recherche assez complexes qui permettent d'obtenir rapidement des résultats spécifiques. Par exemple, on pourra combiner le choix d'un thème (mot-clé) avec certains termes (recherche plein texte) au cours d'une période précise.

### ***e. Autres applications***

#### *Signification des acronymes*

Au grand dam de tous les traducteurs, les auteurs anglophones raffolent des acronymes et abréviations... Le thésaurus de Medline établit automatiquement la correspondance entre certains acronymes et leur forme développée. Ainsi, à la requête portant sur l'acronyme *PTCA*, la base retourne automatiquement : *Percutaneous Transluminal Coronary Angioplasty*. Il faut toutefois être prudent : toutes les significations des acronymes ne sont pas répertoriées. Ainsi, dans Medline, *MLR* signifie *mixed lymphocyte reaction*. Pourtant, le dictionnaire *Medical Abbreviations* (Davis 1999) propose également deux autres solutions : *middle latency response* et *multiple logistic regression*.

Lorsque l'acronyme n'est pas répertorié dans le thésaurus de Medline, ou pour vérifier la possibilité de l'existence d'autres significations, on pourra avoir recours à la recherche plein texte, de nombreux résumés explicitant l'acronyme ou l'abréviation au complet.

#### *Vérification des valeurs numériques et des unités de mesure*

Il arrive que le format des textes soit malmené au cours des transferts électroniques, que certains caractères soient incompatibles entre les différents logiciels de traitement de texte, ou encore que l'auteur se soit trompé... On peut être ainsi amené à vouloir vérifier certains chiffres. Par exemple, devant la phrase : « It is recommended that all American children over the age of two years ingest less than 300 g cholesterol per day. », il y a tout lieu de s'interroger sur la pertinence de la consommation de quantités de cholestérol de l'ordre de 300 g et de se demander s'il s'agit de milligrammes (mg) ou même de microgrammes (µg). Il n'est en effet pas rare que le symbole µ disparaisse purement et simplement ou qu'il soit remplacé par un *m*. Vérifier ces données par une interrogation ciblée permettra de communiquer avec le donneur d'ouvrage avec une efficacité accrue.

#### *Renseignements spécifiques à un médicament*

Notons pour terminer qu'il est possible de rechercher des renseignements spécifiques à un médicament. Toutefois, il faut savoir qu'il faudra interroger les bases, notam-

ment Medline, avec la dénomination commune internationale (DCI) du principe actif (en anglais) et non avec le nom de marque. Les DCI en anglais sont répertoriées dans le thésaurus de Medline.

### Conclusion

Les bases de données bibliographiques électroniques constituent un outil d'une extraordinaire puissance, pour peu qu'on se donne la peine d'en maîtriser le manie- ment. Dans les domaines évoluant rapidement, comme c'est le cas en biomédecine, elles sont un complément indispensable aux manuels de référence, qui sont toujours un peu en retard sur l'état des connaissances, et sur les ouvrages lexicographiques, qui le sont en général bien davantage. Bien qu'elles ne soient pas conçues *a priori* comme outils d'aide à la traduction, les traducteurs, les terminologues comme les chercheurs peuvent cependant tirer parti de leurs nombreuses fonctions d'interroga- tions et ainsi gagner un temps précieux. Enfin, la possibilité d'inclure, de manière raisonnée, de nombreux paramètres essentiels à la portée d'un jugement linguistique nous paraît un avantage majeur sur la plupart des autres ressources offertes sur Internet.

### NOTES

1. On trouvera, datée du début 2000, une liste des revues indexées dans Current Contents, *Clinical Medicine* et *Life Science* mais absentes de Medline (version PubMed) sur le site <<http://www.hospvd.ch/public/chuv/bdfm/ccunique.htm>> (visité le 22 août 2000).
2. Un certain nombre de bases spécialisées, bibliographiques ou non, sont offertes par la *National Library of Medicine* <[http://www.nlm.nih.gov/pubs/factsheets/online\\_databases.html](http://www.nlm.nih.gov/pubs/factsheets/online_databases.html)>.
3. Au moment de mettre sous presse, nous apprenons qu'Internet Grateful Med sera progressivement remplacé par NLM Gateway (<<http://gateway.nlm.nih.gov/gw/Cmd>>), qui permet d'effectuer des recherches dans plusieurs bases de données: Medline/PubMed, OldMedline, LocatorPlus, Aids Meeting, HSR Meetings, HSRProj, MedlinePlus et Dirline (documents informatisés mis en ligne par la National Library of Medicine le 2 février 2001).
4. En fait, les accès gratuits à Medline se multiplient sur Internet, mais les modes d'interrogation, parfois rudimentaires, ne répondent pas nécessairement aux besoins des langagiers. Nous conseillons la consultation de deux sites qui fournissent des informations sur les différents formats offerts: <<http://www.chu-rouen.fr/documed/medline.html>> et <[http://www.ebm.lib.ulg.ac.be/prostate/link\\_med](http://www.ebm.lib.ulg.ac.be/prostate/link_med)>.
5. Fact Sheet – Medical Subject Headings (MeSH), National Library of Medicine of United States, <<http://www.nlm.nih.gov/pubs/factsheets/mesh.html>>, mise à jour du 22 février 2000, visité le 22 août 2000.
6. L'Institut national de la santé et de la recherche médicale (INSERM), en France, a réalisé une version française du thésaurus de Medline, accessible gratuitement: <<http://dicdoc.kb.inserm.fr:2010/basismesh/mesh.html>>. La prudence s'impose toutefois quant à son utilisation en traduction, car il ne s'agit pas d'un dictionnaire bilingue mais d'équivalences de descripteurs dont la structure est parfois étrangement semblable à l'anglais (par exemple, la version « française » du descripteur *myosin light chain kinase* est... *myosin light chain kinase*).
7. « The proteome is the protein complement of a genome. » (Tsugita and Kawakami 1999)
8. « The idea of systematically mapping and identifying the total protein complement of the genome (the "proteome") arose over 20 years ago when the separation of proteins from total cell extracts by two dimensional (2D) gel electrophoresis was developed. »

### RÉFÉRENCES

- BEAUDET, N. et J. HÉBERT (1998): « Medline, mais quel Medline?... L'avis des experts », *Le clinicien*, 13-1, p. 183-186.

- CASSAGNE, H., S. J. DARMONI *et al.* (2000) : *Internet médical professionnel. Guide de bonne pratique clinique et d'enseignement*, Paris, MMI.
- DAVIS, N. (1999) : *Medical Abbreviations*, Huntington Valley, Neil M. Davis Associates.
- HUTCHINSON, D. (1998) : *Medline for Health Professionals. How to Search PubMed on the Internet*, Sacramento (Calif.), New Wind Publishing.
- JAMES, P. (1997) : "Of Genomes and Proteomes," *Biochemical & Biophysical Research Communications*, 231-1, pp. 1-6.
- LODISH, H. *et al.* (2000) : *Molecular Cell Biology*, New York, Freeman & Co.
- TSUGITA, A. and T. KAWAKAMI (1999) : "Proteome Analysis : the State of the Art of the Methodology," *Nippon Hoigaku Zasshi - Japanese Journal of Legal Medicine*, 53-2, pp. 191-198.
- VANDAELE, S. (2001) : « Dérive sémantique et terminologique en traduction médicale », *Congrès international de la FIT (Fédération Internationale des Traducteurs)*, Mons, Belgique, 6-9 août 1999 [à paraître].

## ANNEXE

## 1. Comparaison des versions de Medline offertes par Ovid, PubMed et Grateful Med

	OvidWeb	PubMed	GratefulMed
Correspondance avec les termes de MeSH	Par défaut	Par défaut	Automatique
Interrogation dans les différents champs	Oui, exhaustive	Oui, exhaustive	Limitée ( <i>subject, title word, author name</i> )
Titres dans la langue d'origine	Oui	Non	Oui
Explosion/focus	Oui	Oui	Oui
Accès à l'arborescence	Oui	Oui (Fonction « MeSH browser »)	Oui
Annotations ( <i>Scope Note</i> )	Oui	Oui	Oui
Qualificatifs ( <i>Subheadings</i> )	Oui	Oui	Oui
Historique de la recherche	Oui, complète	Oui, complète	Oui, limitée
Limites	<ul style="list-style-type: none"> <li>- <i>Evidence Based Medicine Reviews</i></li> <li>- Topic Reviews</li> <li>- Cochrane</li> <li>- <i>Article Reviews</i></li> <li>- <i>Best Evidence</i></li> <li>- 4 groupes de périodiques (cancer, nursing, dental)</li> <li>- 13 espèces d'animaux</li> <li>- 68 langues (dont le français)</li> <li>- 9 groupes d'âge</li> <li>- 48 types d'article</li> <li>- groupes d'études (humains/animaux)</li> <li>- sexe (fém./masc.)</li> <li>- années 1966-2000</li> </ul>	<ul style="list-style-type: none"> <li>- 7 types de publications</li> <li>- 9 groupes d'âge</li> <li>- date d'inscription</li> <li>- 7 langues, dont le français</li> <li>- 9 bases</li> </ul>	<ul style="list-style-type: none"> <li>-7 langues dont le français</li> <li>-7 types de publications</li> <li>- groupes d'études (humains/animaux)</li> <li>- sexe (fém./masc.)</li> <li>- 9 groupes d'âges</li> <li>- années 1966-2000</li> </ul>

Combinaisons (opérateurs booléens)	Oui	Oui	Oui
Sauvegarde de documents	Nombreuses options de sauvegarde	Oui	Oui
Remarques	Interface d'interrogation conviviale et très complète		

## 2. Adresses Internet

### Current Contents

<<http://ovid/gwdg/de/ovidweb/fldguide/cc.htm>>

<<http://www.isinet.com/isi/products>>

### Cab Abstracts

<http://www.cabi.org/infolib/cababdb/cababdb.htm>

### Pascal

<<http://www.inist.fr/pascal/pascal.htm>>

### Article@INIST

<<http://form.inist.fr/public/fre/conslt.htm>>

### HealthGate

<<http://www.healthgate.com>>

### Ovid

<<http://www.ovid.com>>

### ICIST

<[http://www.cnrc.ca/cisti/irm/statit\\_f.shtml](http://www.cnrc.ca/cisti/irm/statit_f.shtml)>

Accès aux usagers de certaines bibliothèques, par exemple, de celles de l'Université de Montréal:

<<http://www.bib.umontreal.ca/SB/swets.htm>>

### Répertoire de bases de données francophones et anglophones

<<http://www.chu-rouen.fr/documed/tex.html>>

## 3. Sites d'accès à Medline

### National Library of Medicine (liste des bases avec leurs caractéristiques)

<[http://www.nlm.nih.gov/pubs/factsheets/online\\_databases.html](http://www.nlm.nih.gov/pubs/factsheets/online_databases.html)>

### Medline: Internet Grateful Med v. 2.6.3

<<http://igm.nlm.nih.gov>>

### Medline: PubMed

<<http://www.ncbi.nlm.nih.gov/80/entrez>>

### Medline: Ovid (accès réservé)

<<http://hermes.bib.umontreal.ca/ovidweb>> (Université de Montréal)

### Medline: liste de sites d'accès (description du contenu et caractéristiques)

<[http://www.ebm.lib.ulg.ac.be/prostate/link\\_med.htm](http://www.ebm.lib.ulg.ac.be/prostate/link_med.htm)> (Université de Liège – dernière mise à jour 21 août 2000)

### National Library of Medicine: bases MEDLARS

<[http://www.nlm.nih.gov/pubs/factsheets/online\\_databases.html](http://www.nlm.nih.gov/pubs/factsheets/online_databases.html)>

## 4. Modes d'emplois et informations

### Medline (notions générales et utilisation de la version PubMed)

<<http://www.urfist.jussieu.fr/urfist/biolo/bioguide2/medline/intro1.htm>>

**Medline (version PubMed)**

<[http://bu2.timone.univ-mrs.fr/guide\\_pubmed.asp](http://bu2.timone.univ-mrs.fr/guide_pubmed.asp)>

**Medline (version Ovid/Web)**

<<http://www.chu-rouen.fr/documed/mme.html>>

Consulter également Hutchinson (1998).

**Document rassemblant des informations sur Medline (accès, manuels d'utilisation, études comparatives et commentaires ; accès à d'autres bases)**

<<http://www.chu-rouen.fr/documed/medline.htm>>

**Analyse comparative des différents Medline offerts sur Internet**

Beaudet et Hébert (1998)

**REMERCIEMENTS**

Le présent travail a pu être réalisé grâce à des subventions octroyées par l'Université de Montréal. L'auteure remercie Manon Pageau et Micheline Roberge pour l'aide apportée lors de la collecte des données.